CLAIMS

Having thus described our invention, what we claim as new and desire to secure by Letters Patent is as follows:

1. A predictive model method, comprising:

receiving an input data into an initial model to develop an initial model output; and

receiving both of said input data and said initial model output as input data into a first transform/regression stage.

2. The method of claim 1, further comprising:

providing an output of said first transform/regression stage as a first of two inputs into a second transform/regression stage,

wherein a second of said two inputs comprises said input data into said initial model.

3. The method of claim 2, further comprising:

successively providing, for one or more additional stages, an output of a preceding transform/regression stage as a first of two inputs into a next transform/regression stage,

wherein a second of said two inputs comprises said input data into said initial model.

4. The method of claim 1, wherein said first transform/regression stage comprises:

a feature transform stage receiving said input data and said initial model output;

a linear regression stage receiving an output of said feature transform stage; and

an output summing node receiving as inputs said initial model output and an output of linear regression stage, an output of said output summing node comprising a first stage model output.

5. The method of claim 4, further comprising:

successively providing, for one or more stages, an output of a preceding transform/regression stage as a first of two inputs into a next transform/regression stage,

wherein a second of said two inputs comprises said input data into said initial model.

6. The method of claim 5, wherein for each said one or more stages, a third input into said next transform/regression stage comprises an output of said linear regression stage of said preceding transform/regression stage, and, for each said transform/regression stage, an output of said linear regression stage for said linear regression stage is carried forward to be an input into all successive transform/regression stages.

7. The method of claim 5, further comprising:

avoiding an overfitting in said predictive model by determining when a successive stage does not add to a performance of said predictive model.

8. The method of claim 7, wherein said determining of performance degradation comprises a holdout method, said holdout method comprising:

dividing an available data into a training set and a holdout data set;

using said training set to estimate a model parameter and to construct an alternative model structure; and

using said holdout data set to make a selection among said alternative model structure.

9. The method of claim 7, wherein said determining of performance degradation comprises a cross-validation method, said cross-validation method comprising:

dividing an available data into a plurality of folds of data; and

successively, using each said fold as a holdout data set, and a remaining data not in said fold is used as a training data set to estimate model parameters and to construct alternative model structures and said training data set is used to make a selection among said alternative model structures.

10. A predictive modeling method, comprising:

establishing an initial model module to instance an initial model; and

establishing a stage model module to instance a stage model for each of a plurality of successive stages,

wherein each instanced stage model receives, as an input, an output from a preceding stage model, and each said instanced stage model provides a stage model output that is used as an input into any next successive instanced stage model.

11. The method of claim 10, wherein each said stage model feeds forward a second output as another input into all succeeding stage models..

12. The method of claim 10, further comprising:

instancing said initial model;

instancing a first stage model that receives an output data from said initial model as an input data and to provide a first stage model output;

successively instancing one or more of said stage models to be successive stage models, wherein a first successive stage model receives said first stage output data as an input data and provides an output data to be an input data to a second successive stage model, if any, and each successive stage model, if any, receives a stage output data from an immediately preceding successive stage mode and each successive stage model provides a stage output data to become an input data to a next successive stage model; and

providing an input data as inputs to said initial model, said first stage, and each said successive stage model.

13. The method of claim 12, further comprising:

determining when an additional successive stage would not add to a performance of the predictive model.

14. The method of claim 13, wherein said determining of performance degradation comprises a holdout method, said holdout method comprising:

dividing an available data into a training set and a holdout data set;

using said training set to estimate a model parameter and to construct alternative model structures; and

using said holdout data set to make a selection among said alternative model structures.

15. The method of claim 13, wherein said determining of performance degradation comprises a cross-validation method, said cross-validation method comprising:

dividing an available data into a plurality of folds of data;

successively, using each said fold as a holdout data set, and a remaining data not in said fold is used as a training data set to estimate model parameters and to construct alternative model structures and said training data set is used to make a selection among said alternative model structures.

16. The method of claim 10, wherein said stage model comprises:

a first data input port;

a second data input port;

a feature transform stage receiving data from said first data input port and said second data input port;

a linear regression stage receiving an output from said feature transform stage;

a summing node receiving data from said first data input port and output data from said linear regression stage; and

an output port receiving data outputted from said summing node.

17. The method of claim 16, wherein said stage model further comprises:

a second output port to provide said output data from said linear regression stage to be a second output from said stage model; and

one or more input ports to receive data from said second output port of preceding stages to be input data into said feature transform stage.

18. An apparatus to perform a predictive modeling method, said apparatus comprising:

an initial model module to instance an initial model; and

a stage model module to instance a stage model for each of a plurality of successive stages,

wherein each said stage model receives an input from an immediately preceding stage and provides an output to a next succeeding stage.

19. The apparatus of claim 18, further comprising:

a controller to cause said initial model and each of a plurality of said successive stage models to be instanced and to interconnect said initial model and said plurality of successive stage models; and

a graphic user interface to allow a user to control said controller and said predictive modeling method, to input data into said initial model, and to one of display and print to one of a printer, a data file, and an application program the output of a final one of said successive stage models.

20. A signal-bearing medium tangibly embodying a program of machine-readable instructions executable by a digital processing apparatus to perform a predictive modeling method, said instructions comprising:

an initial model module to instance an initial model; and

a stage model module to instance a stage model for each of a plurality of successive stages,

wherein each instanced stage model receives, as an input, an output from a preceding stage model, and each said instanced stage model provides a stage model output that is used as an input into a next successive instanced stage model.

21. The signal-bearing medium of claim 20, wherein said instructions further comprise:

causing said initial model and each of a plurality of said successive stage models to be instanced and to appropriately interconnect said initial model and said plurality of successive stage models;

allowing a user to control said controller and said predictive modeling method, to input data into said initial model, and to one of display and print to one of a printer, a data file, and an application program the output of a final one of said successive stage models;

receiving input data; and

allowing an output data of said predictive modeling method to be provided as output data.

22. A method of providing a service, said method comprising at least one of:

providing an execution of a predictive modeling method, wherein said predictive modeling method comprises:

establishing an initial model module to instance an initial model; and

establishing a stage model module to instance a stage model for each of a plurality of successive stages, wherein each instanced stage model receives, as an input, an output from a preceding stage model, and each said instanced stage model provides a stage model output that is used as an input into a next successive instanced stage model.

23. A predictive modeling method, comprising:

using an initial model that provides an initial model of input data; and

using at least one successive stage model, each said successive stage model providing a cross-product interaction model.

24. A method of determining performance degradation in an iterative predictive modeling, said method comprising:

dividing an available data into a training set and a holdout data set;

using said training set to estimate a model parameter and to construct alternative model structures; and

using said holdout data set to make a selection among said alternative model structures.

25. A method of determining performance degradation in an iterative predictive modeling, said method comprising:

dividing an available data into a plurality of folds of data;

successively, using each said fold as a holdout data set, and a remaining data not in said fold is used as a training data set to estimate model parameters and to construct alternative model structures and said training data set is used to make a selection among said alternative model structures.

26. A method for deploying computing infrastructure, comprising integrating computer-readable code into a computing system, wherein the code in combination with the computing system is capable of performing the method of claim 1.